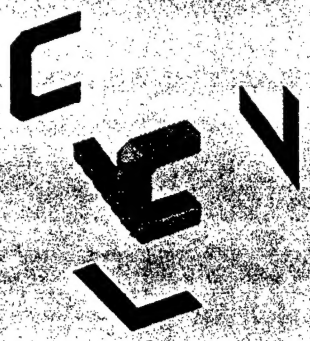
CAR-TR-885
CS-TR-3898

N00014-95-0521
April 1998

Data-driven Multi-channel Super-resolution with Application to Video Sequences*

H. Shekarforoush and R. Chellappa
Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3275

COMPUTER VISION LABORATORY



CENTER FOR AUTOMATION RESEARCH

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742-3275

CAR-TR-885
CS-TR-3898

N00014-95-0521
April 1998

Data-driven Multi-channel Super-resolution with Application to Video Sequences*

H. Shekarforoush and R. Chellappa
Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3275

Abstract

A method is proposed for super-resolving multi-channel data with applications to PREDATOR video sequences. Using a generalization of Papoulis' sampling theorem, a closed-form solution has been obtained leading to a high-speed algorithm which can be realistically applied to large data sets such as video sequences. In existing multi-frame methods it is a common practice to assume that the channel transfer functions are known and invariant from one frame to another, using empirical models such as Gaussian, sinc, etc. We have assumed that the transfer functions are unknown and may vary even when the same sensor is employed, and hence use the observed data to derive the Point Spread Function (PSF) for each frame. The estimated PSFs are used in the super-resolution algorithm. Results on PREDATOR video images are then given.

Keywords: Super-resolution, video data processing, multi-channel sampling, data-driven estimation

This work was partially supported by ONR Grant N00014-95-1-0521.

1 Introduction

Super-resolution refers to methods that attempt to increase resolving power by means of image processing. Super-resolving video data in particular is potentially of significant interest in visual surveillance and monitoring of human and vehicular activities in both civilian and battlefield applications [10][23]. In fact, many automatic target recognition, detection and identification problems suffer from lack of adequate resolution of the data (PREDATOR video is a good example). When replacing a sensor is not a feasible solution (due to either cost or technical limitations), post-processing by means of super-resolution algorithms is the only means of improving the resolving power for enhancing performance.

Earlier work on super-resolution involved single-frame methods in which, assuming that the signal was of compact support in space (time), its spectrum was extrapolated outside a known support. Variations of single-frame methods include: methods based on analytic continuation [5], those based on prolate spheroidal wave functions [20][21][22], the well-known Papoulis-Gerchberg [3][12] algorithm, and some probabilistic methods [7].

Recent work has mostly involved multi-frame data [1][6][8][9][17][19], where the assumption is that the signal is not space-limited but rather band-limited and that each frame has been down-sampled so that every frame is an aliased representation of the underlying signal. By making the assumption that the sampling matrices are not correlated one can then attempt to increase the sampling density by trading off the temporal bandwidth, and hence to unscramble some or all of the aliased portion of the spectrum.

In this paper we will be mainly concerned with multi-frame methods, which clearly are better adapted to processing video data. However, existing methods suffer from a number of shortcomings which we have addressed:

- In the existing literature, the sensor PSF is assumed invariant from one frame to another; empirical models such as Gaussian, sinc, etc. are generally used. The analysis is thus simplified at the cost of reduced performance. We propose a method where the PSF is adaptively estimated from the observed data for each frame in the sequence. By allowing different PSFs for different frames, our method extends multi-frame methods to a true multi-channel framework.

- Speed is an important issue when dealing with video sequences. Most existing methods are slow. We have proposed a closed-form solution, resulting in a high-speed algorithm and allowing for realistic application to video data.

2 Multi-channel super-resolution

Our problem is formulated as follows: We have a sequence of M images that were collected by a set of systems whose transfer functions are unknown. The output of each channel is down-sampled, in the sense that each image alone cannot represent the underlying input signal. We are interested in establishing a method of reconstructing the input signal, as well as conditions under which the reconstruction leads to a physically meaningful solution.

Our analysis is based on a generalization of Papoulis' sampling theorem [13]. Consider the very simple example of reconstructing a 1D signal observed by two identical sensors sampling uniformly with the same period T but with a temporal shift of δt between them. If δt is an integer multiple of T then the two sensors will produce redundant information, using twice as much temporal bandwidth as required. Therefore, in order to be able to trade off some of the temporal bandwidth, we must have $\delta t \neq kT \quad \forall k \in Z$. In general, this sampling is non-uniform unless $\delta t = (k + \frac{1}{2})T$ for some integer k. However, the samples constitute a special case of non-uniform sampling referred to as non-uniform recurring samples, bunched samples or interlaced samples. This is shown in Figure 1.

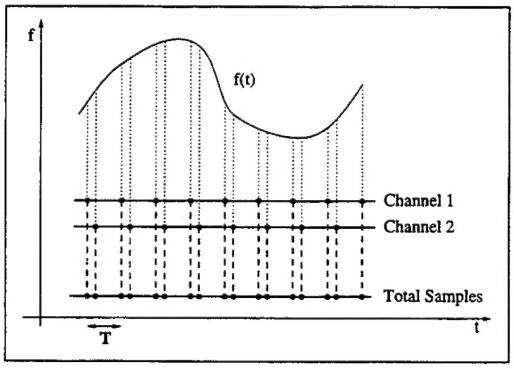


Figure 1: Multi-channel interlaced sampling

Therefore, even in the simplest case, the problem is outside the scope of uniform sampling theory. In fact, in practice the problem can be much more complicated due to the following factors:

- In higher dimensions inter-frame shifts will generalize to other transformations whose parameters are in general unknown and need to be accurately estimated (the images need to be registered at sub-pixel accuracy).
- Sensors in general cause degradations due to their PSFs.

- Sensors may be different (have different PSFs), or in the case when the same sensor is used, the PSF for each may vary frame due to lighting conditions, noise, etc.

Below, we will first set the theoretical background for designing our algorithm.

2.1 Theory

Definition 2.1 *A set $\mathcal{U}(C) \subset \mathbb{R}^n$ is said to be a set of uniqueness for a given class of functions C if $f = 0$ on $\mathcal{U}(C)$ implies that f is the null function: $f \equiv 0$.*

In general, a set of uniqueness does not have to be discrete, but all sets of uniqueness considered here will be discrete subsets of $\mathbb{R}^n$, i.e. multivariate sequences. Also the functions we will be dealing with will belong to the class of entire functions of exponential type, namely functions that belong to a Paley-Wiener space $\mathcal{P}$ [11] over a compact support $\mathbf{B} \subset \mathbb{R}^n$. The following theorem will then generalize Papoulis' sampling theorem to a multi-dimensional multi-resolution framework, where multi-resolution implies that, contrary to Papoulis' theorem, the channel sampling densities do not have to be equal, provided that the overall density equals that of Nyquist.

Theorem 2.1 *(Multi-Resolution, Multi-channel Band-limited Sampling) Let $f(\mathbf{t}) \in \mathcal{P}(\mathbf{B})$ be convolved with a set of mutually uncorrelated linear shift-invariant filters in $\ell_2(\mathbb{R}^n)$, whose spectra $\{\hat{h}_m(\mathbf{u})\}_{m=1,\ldots,M}$ are uniformly bounded away from zero in $\mathbf{B}$. Assume also that the outputs are known only on a discrete set of points $\{g_m(\mathbf{S}_m\mathbf{k})\}_{m=1,\ldots,M}$, where the nonsingular matrices $\mathbf{S}_m$ are such that*

$$\bigcup_{m=1}^{M} \{\mathbf{S}_m\mathbf{k}\}_{\mathbf{k}\in Z^n} = \mathcal{U}\left(\mathcal{P}(\mathbf{B})\right) \tag{1}$$

Moreover, let $\{\varphi_m(\mathbf{t})\}_{m=1,\ldots,M}$ be a set of linear shift-invariant functions in $\ell_2(\mathbb{R}^n)$ satisfying the following biorthogonality conditions over $\mathbf{B}$:

$$\hat{h}_m(\mathbf{u})\hat{\varphi}_m(\mathbf{u}) = 1_{\mathbf{B}} \tag{2}$$

Then there exist scalars $c_{m\mathbf{k}}$ such that

$$f(\mathbf{t}) = \frac{1}{M}\sum_{m=1}^{M}\sum_{k\in Z^n} c_{m\mathbf{k}}\varphi_m(\mathbf{t} - \mathbf{S}_m\mathbf{k}) \tag{3}$$

with convergence being uniform and $c_{m\mathbf{k}} = g_m(\mathbf{S}_m\mathbf{k})$.

For a proof see [16].

In fact, $\{\varphi_m(\mathbf{t} - \mathbf{S}_m\mathbf{k})\}_{m=1,\ldots,M}$, which we shall refer to as the biorthogonal set, constitutes a Riesz basis in $\mathcal{P}(\mathbf{B})$, and each component in the set can be found as a space-varying linear transformation of Papoulis' periodic inverse filters (see [16] for details).

In contrast to Papoulis' theorem, where the reconstruction algorithm is based on solving a large linear matrix equation, the above theorem provides a more straightforward method. In fact, note that the biorthogonal set is known if the channel transfer functions (i.e., the PSFs) can be estimated and the sampling matrices can be specified. In image processing terminology, the latter implies that the image frames need to be registered. For the registration we have used a method based on phase correlation developed in [2][18].

As for the reconstruction, in principle, any basis can be used (wavelet, orthogonal polynomials, etc.). The Fourier basis, for instance, can be found using the duality due to the Paley-Wiener theorem [11]:

$$\{\hat{\varphi}_m(\mathbf{u})exp(-i\langle\mathbf{S}_m\mathbf{k}, \mathbf{u}\rangle)\}_{m=1,\ldots,M} \tag{4}$$

where, for our application, the $\mathbf{k}$'s take their values in compact subsets of Z^2, $\langle\rangle$ is the usual inner product in $\mathbb{R}^2$, and $\mathbf{u}$ (with abuse of notation) will hereafter stand for discrete frequency vectors sampled at super-resolution density. Note from equation (4) that the expansion of each frame on the biorthogonal basis can be found by expanding on the standard orthonormal basis followed by applying the biorthogonal filter $\hat{\varphi}_m(\mathbf{u})$ (a projection). The basis is completely specified if we can find $\hat{\varphi}_m$'s which are in turn specified according to the biorthogonality condition in terms of the channel transfer functions. Once the channel transfer functions have been estimated, we can reconstruct the biorthogonal set (or a pseudo-biorthogonal set if the transfer functions contain singularities). Therefore, in the next section we will develop a data-driven method for estimating the PSFs.

2.2 Estimating the PSFs

In this section we will define our PSF model, or more precisely its spectrum, refered to as the Optical Transfer Function (OTF). We will then develop a method for data-driven estimation of the OTFs. As we will see in the next section, the proposed method is particularly applicable to video data, since it requires two images with only relative shifts between them. In fact, in video sequences (such as PREDATOR data) transformations between successive frames can be closely approximated by shifts and a small-angle rotation within the image plane. Therefore, by using a registration method (see the section on implementation), successive frames

are rotation-compensated prior to PSF estimation.

In the existing literature on super-resolution, it is a common practice to assume that the OTF is known and traditionally empirical models such as Gaussian, sinc, etc. are adopted. Although these models can greatly reduce the complexity in both analysis and implementation due to their nice global behavior (isotropy around the origin), they fail to capture local variations of the OTF and hence are not useful for adaptive estimation and inversion.

We will show below how more versatile local adaptive models can be built using local isotropy. For this purpose we will assume that the OTFs of the imaging systems are linear shift-invariant and of finite duration and hence can be represented by FIR filters:

$$\hat{h}_m(\mathbf{u}) = \sum_{\mathbf{k}\in\mathbf{T}} c_k \exp(-i\langle \mathbf{u}, \mathbf{k}\rangle) \tag{5}$$

where the c_ks are constants, and $\mathbf{T}$ is a compact subset of Z^2.

As is well known, due to the fundamental theorem of algebra, in the univariate case the transfer function $\hat{h}_m$ can be factored as the product of its roots:

$$\hat{h}_m(u) = A\prod_{k=0}^{N}(1 - a_k \exp(-iu)) \tag{6}$$

$$= A\prod_{k=0}^{N}(1 - a_k z^{-1}) \tag{7}$$

where N specifies the extent of the impulse response, the a_ks and A are constants, and $z = \exp(iu)$.

Clearly $\hat{h}_m$ is then specified up to a scale factor if its roots are known. Unfortunately, due to the general impossibility of factoring polynomials in higher dimensions, this simple convenient factorization of the spectrum in univariate problems does not extend to $n > 1$ dimensions.

However, we will show that in the two-dimensional case the problem can have a solution too. We start by assuming that the zeros of our channel transfer functions occur at isolated points[1]. Since the transfer functions are then either locally convex or concave around their zeros, one may assume local isotropy in small neighborhoods of the roots. The following results will show how a two-dimensional spectrum with isolated zeros can be factored.

[1]In general, the zeros of an entire function of two or more variables do not have to be isolated. However, in practice they often occur at isolated points [15].

Proposition 2.1 *Let $\{\zeta_n\}_{n\in Z}$ be a set of isolated points in $\mathbb{C}^n$. Then any function defined over $\mathbb{C}^n$, which is locally isotropic around these points, can be decomposed into a sum of functions, each globally isotropic around one point of the set.*

The existence of such a decomposition under appropriate boundary conditions is obvious. However, the decomposition is not necessarily unique, which suggests that the inverse may not necessarily be true. In other words, a sum of functions that are globally isotropic around distinct isolated points will not necessarily yield a function that is locally isotropic around these points. In fact, counter-examples can be readily found.

Theorem 2.2 *Let $\hat{h}_k(u,v)$ be the spectrum of a band-limited function vanishing outside a compact support $\mathbf{T} \subset \mathbb{R}^2$. Let also $\hat{h}_k(u,v)$ be isotropic around the point (u_k, v_k). Then*

$$\lim_{b_k \to 0} b_k \hat{h}_k(u,v) = exp(-i\omega t_k) \tag{8}$$

where $\omega = ((u-u_k)^2 + (v-v_k)^2)^{\frac{1}{2}}$ and
$t_k = (x_k^2 + y_k^2)^{\frac{1}{2}}$ with $(x_k, y_k) \in \mathbf{T}$.

For a proof see the appendix.

Let us now assume that the channel transfer functions are locally isotropic around their roots. Then according to Proposition 2.1 we can write

$$\hat{h}_m(u,v) = \sum_k \hat{h}_{mk}(u,v) \tag{9}$$

where for every m, $\hat{h}_{mk}$ is globally isotropic around one root of $\hat{h}_m$. We shall denote the roots of $\hat{h}_m$ by the set $\{(u_k, v_k)\}_{k\in K}$, where K is a compact index set.

Therefore, an OTF which is locally isotropic around its roots can be expanded as follows:

$$\hat{h}(u,v) = \sum_k a_k exp(-i\omega t_k) \quad where \;\; a_k = b_k^{-1} \tag{10}$$

To physically interpret this convergence at the limit, notice that the b_ks define the radii of local isotropy around the roots. Therefore, the more local the isotropy is, the better the OTF can be approximated by an expansion of the form (10). This assumption of local isotropy clearly relaxes the severe constraints usually imposed in the literature by global symmetries.

2.3 Spectral Factorization

In the classical literature this term is used when it is required to find a function whose power spectrum is known. The problem has a solution if the power spectrum satisfies

the Paley-Wiener condition [14]. Below we will use the term in a more general context where two functions are sought whose cross power spectrum is known. We will show that the solution can be found if the functions are of compact support (e.g. FIR filters). The support constraint condition is clearly equivalent to the Paley-Wiener condition due to the uncertainty principle [14].

From the expansion in (10) and the fundamental theorem of algebra, it immediately follows that

$$\hat{h}(u,v) = A\prod_k (1 - c_k z^{-1}) \tag{11}$$

where A and the c_ks are constants and $z = \exp(i\omega) = \exp\left(i\left((u-u_k)^2 + (v-v_k)^2\right)^{\frac{1}{2}}\right)$.

Clearly, the factorization is only valid if the c_ks are identically equal to unity. Therefore, if the roots of the OTF are specified, the OTF is known up to a scale factor. We will deal with the scale factor shortly. Let us first see how the roots of the OTF can be estimated from two observations of the same scene.

Consider the case where an object function has been observed by two systems modeled as

$$\hat{g}_1 = \hat{h}_1 \hat{f}_1 + \hat{n}_1 \tag{12}$$
$$\hat{g}_2 = \hat{h}_2 \hat{f}_2 + \hat{n}_2 \tag{13}$$

where f_2 is a shifted version of f_1, $\hat{h}_1$ and $\hat{h}_2$ are the OTFs, $\hat{n}_1$ and $\hat{n}_2$ are random noise processes, and $\hat{g}_1$ and $\hat{g}_2$ are the observed image spectra.

When $\hat{f}_1$ and $\hat{f}_2$ are uniformly bounded away from zero on their supports, we can write

$$\hat{g}_1 = \hat{h}_{g1} \hat{f}_1 \tag{14}$$
$$\hat{g}_2 = \hat{h}_{g2} \hat{f}_2 \tag{15}$$

where $\hat{h}_{g1} = \hat{h}_1 + \hat{n}_1 \hat{f}_1^*$ and $\hat{h}_{g2} = \hat{h}_2 + \hat{n}_2 \hat{f}_2^*$, with * denoting the complex conjugate. We shall refer to such transfer functions as the Generalized Transfer Functions (GTFs) of the imaging systems.

We can now note that the magnitude of the cross power spectrum is given by

$$\mid \hat{g}_1 \hat{g}_2^* \mid = \mid \hat{h}_{g1} \hat{h}_{g2}^* \mid \tag{16}$$

Therefore, since both $\hat{h}_{g1}$ and $\hat{h}_{g2}$ are assumed to be FIR filters, the poles of the cross power spectrum will be uniquely determined by the roots of $\hat{h}_{g2}$. In practice this implies that the roots of $\hat{h}_{g2}$ will be present in the form of singularities, appearing

as a set of spikes scaled at different frequencies. In other words, the roots can be identified by simple inspection of the magnitude of the cross power spectrum. Figure 2 shows an example of these spike patterns.

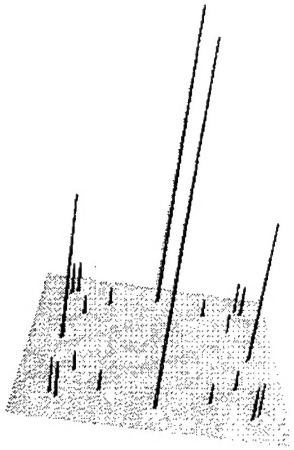


Figure 2: Typical spike patterns appearing in the magnitude of a cross power spectrum corresponding to isolated zeros of one of the transfer functions

To determine the scale value A we will assume that the GTFs preserve the mean value of the object function and hence

$$A = \frac{1}{\prod_k (1 - exp(-i\gamma_k))} \tag{17}$$

where $\gamma_k = (u_k^2 + v_k^2)^{\frac{1}{2}}$.

Therefore, by identifying the singularities (spikes) in the magnitude of the cross power spectrum, we can completely specify the corresponding GTF.

2.4 Constructing Pseudo-biorthogonals

In the absence of noise the GTF is equal to the OTF and when the OTF is an ideal all-pass filter it corresponds to the noise process. We shall now use a simple method for constructing the biorthogonal set. In fact, since in our model the GTFs (or the OTFs) are assumed to have a set of isolated zeros, we can only attempt to construct a pseudo-biorthogonal set. One may consider several approaches for constructing a stable pseudo-biorthogonal set, for instance by using an approximation around the origin.

Below, we will use a simple method which is based on the following identity derived from the biorthogonality condition:

$$(1 + \hat{\varphi}_m)^{-1} \equiv \hat{h}_m (1 + \hat{h}_m)^{-1} \tag{18}$$

Using a first-order expansion of the left-hand side around the origin, we find that

$$1 - \hat{\varphi}_m + \mathcal{O}(\hat{\varphi}_m^2) \simeq \hat{h}_m (1 + \hat{h}_m)^{-1} \tag{19}$$

and hence

$$\hat{\varphi}_m \simeq \frac{\hat{h}_m^*}{\hat{h}_m \hat{h}_m^* + \hat{h}_m^*} \tag{20}$$

which has some resemblance to the standard Wiener filter.

3 Implementation and Results

We have implemented our method using a Fourier basis. The following are the steps in the algorithm:

- For a sequence of M successive frames the PSFs are estimated by estimating the inter-frame rotations [2] and then using the method described above.
- The pseudo-biorthogonals are constructed for projection onto optimal bases.
- Each frame is decomposed on a standard orthonormal Fourier basis and then projected onto a more optimal Riesz basis using the corresponding pseudo-biorthogonal transfer function.
- Using sub-pixel inter-frame displacement values [18], successive frames are re-aligned and added into a single frame according to (3).

The method has been tested on numerous images from PREDATOR video sequences, some of which are shown below. The results have been compared to interpolation by Shannon's sampling expansion (i.e. infinite-order interpolation) followed by image sharpening.

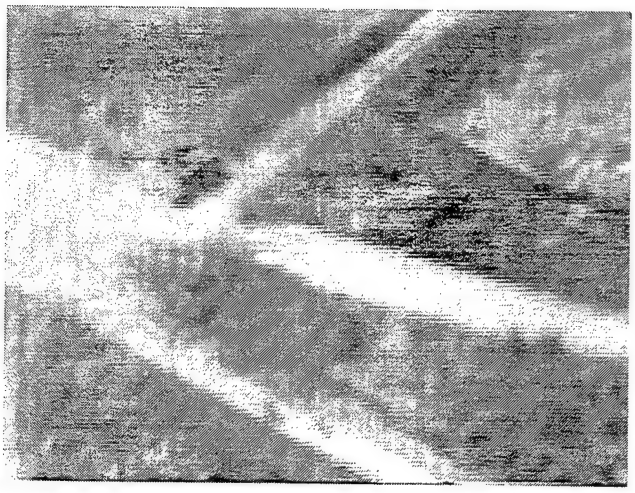
In each sequence presented below, image (a) shows one of the low-resolution frames in a sequence of images; image (b) is its interpolated version; and image (c) is the super-resolved version. Interpolating using Shannon's sampling expansion is, in fact, equivalent to decomposing the frame on the standard orthonormal basis and then ideally low-pass filtering for reconstruction on a denser basis. Clearly, lower-order interpolations (such as bilinear) would yield even lower quality images. The interpolations in (b) have two main shortcomings compared to the super-resolved images in (c):

- As is well known, an orthonormal basis is only sub-optimal in the presence of noise and other degradations and hence in practice Shannon's interpolation can only provide sub-optimal results. Although, when the signal to noise ratio is high, interpolation followed by sharpening may result in good results, for a very low signal to noise ratio this approach can considerably enhance the noise and the artifacts of the contaminated signal. Therefore, an advantage of our super-resolution algorithm is its capability of resolving visual information in noisy data such as PREDATOR video.

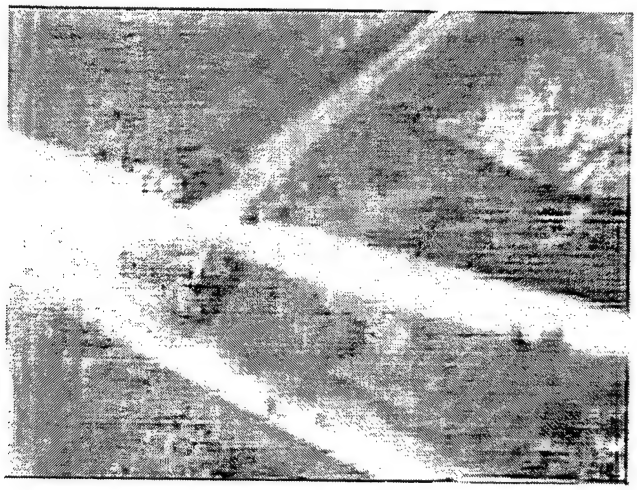
- When using Shannon's interpolation, samples are only specified by a single frame and hence the temporal bandwidth is not exploited, as in our super-resolution method, for increasing the sampling density.

Note that the images used for experimentation are real data for which the ground truth is not available and therefore it is not possible to investigate the quality of the super-resolved or interpolated images in terms of signal-to-noise ratio or other quantifying measures. However, our future plan is to investigate the quality of the results in terms of the performance improvement in state-of-the art automatic target recognition, detection and identification methods. Results should be available in a comprehensive report in the near future.

(a)

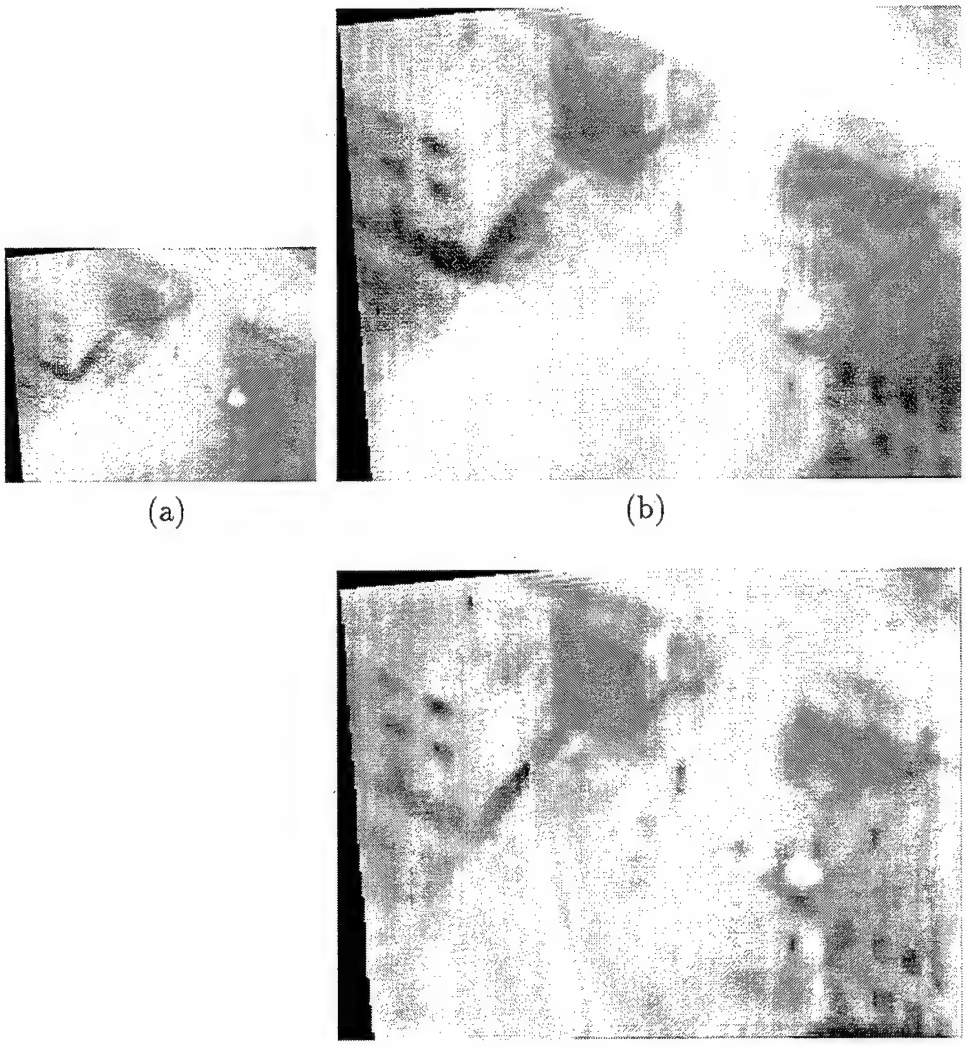


(b)



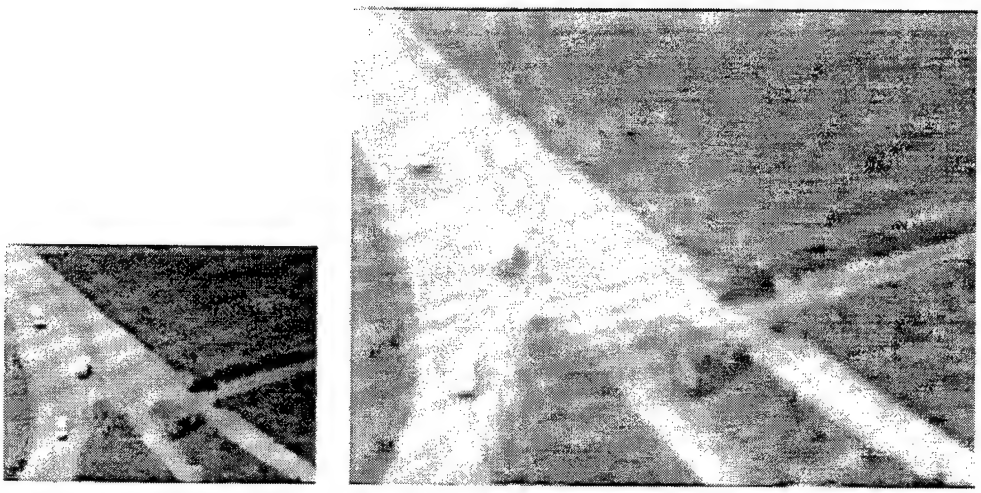
(c)

Figure 3: (a) One of the test images in the sequence, (b) interpolation and sharpening, (c) super-resolved image



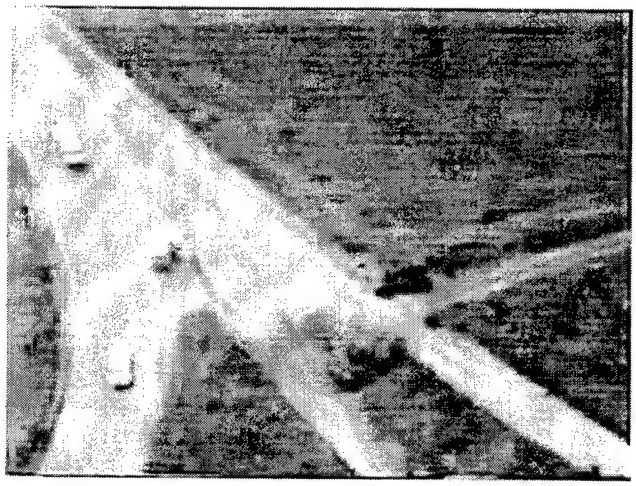
(a) (b) (c)

Figure 4: (a) One of the test images in the sequence, (b) interpolation and sharpening, (c) super-resolved image



(a)

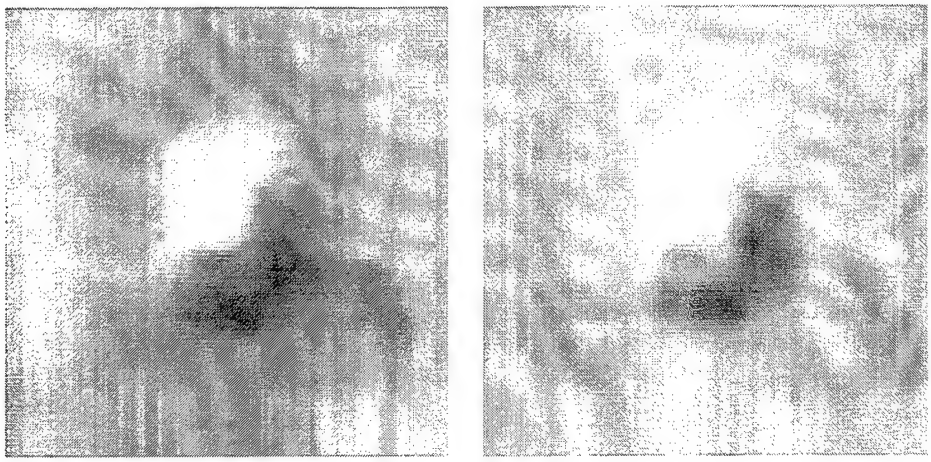
(b)



(c)

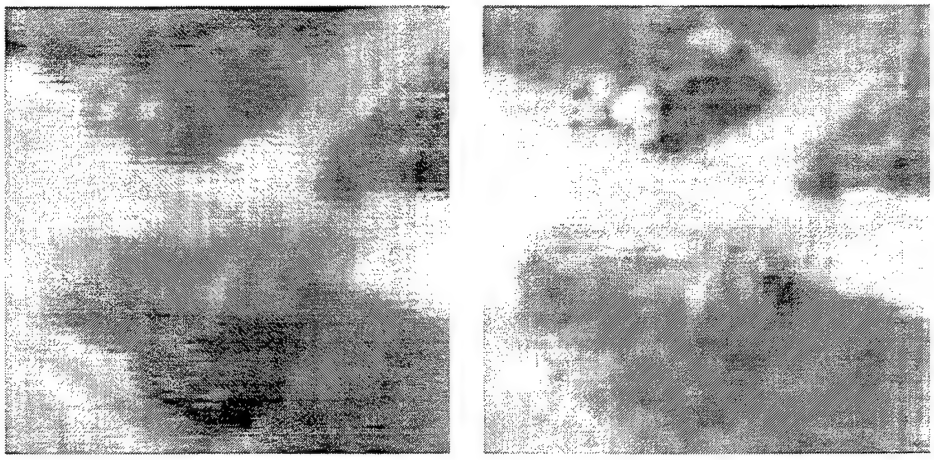
Figure 5: (a) One of the test images in the sequence, (b) interpolation and sharpening, (c) super-resolved image

In the following images we have zoomed in to some regions in Figure 5 and Figure 3 where a vehicle and a tank can be seen closely in both the interpolated/sharpened version and the super-resolved version. One can easily see the clear recovery of the edges of the vehicle and tank in the super-resolved image compared to the artifactual versions obtained by interpolation and sharpenning.



(a) (b)

Figure 6: Zoomed areas: (a) Interpolated/sharpened, (b) super-resolved



(a) (b)

Figure 7: Zoomed areas: (a) Interpolated/sharpened, (b) super-resolved

4 Conclusion

A method of multi-channel super-resolution has been proposed, based on projecting image frames in a sequence onto a more optimal Riesz basis and exploiting the temporal bandwidth to increase the spatial resolution. Several contributions have been made:

- A unified mathematical framework based on non-uniform sampling theory.
- Direct extraction of channel transfer functions from the input data, allowing for the design and implementation of sensor-dependent reconstruction and adaptive estimation of variations in the imaging environment, e.g. the PSF and the noise.
- A closed-form solution leading to a high speed algorithm: typically 15 seconds per 256×256 frame, using 10-channel data fusion, on an Ultra Sparc 1.

Thus the multi-frame setup commonly used in the literature has been extended to a multi-channel setup, where the imaging parameters are assumed to vary from one frame to another. This generalization also allows extension of the method to other sensor types such as radar or infra-red.

References

[1] M. Berthod, H. Shekarforoush, M. Werman, and J. Zerubia. Reconstruction of high-resolution 3D visual information. In *Proc. CVPR*, pages 654–657, 1994.

[2] E. DeCastro and C. Morandi. Registration of translated and rotated images using finite Fourier transforms. *IEEE Trans. PAMI*, 9:700–703, 1987.

[3] R. W Gerchberg. Super-resolution through error energy reduction. *Optica Acta*, 21:709–720, 1974.

[4] A. Gray and T. M. Mathews. *A Treatise on Bessel Functions and Their Applications to Physics.* Dover Publications, 1966.

[5] J. L. Harris. Diffraction and resolving power. *J. Opt. Soc. Am.*, 54:931–936, 1964.

[6] T. S. Huang and R. Y. Tsai. *Image Sequence Analysis: Motion Estimation*, chapter 1, pages 1–18. Springer-Verlag, 1981.

[7] B. R. Hunt. Imagery super-resolution: Emerging prospects. *SPIE Applications of Digital Image Processing*, 1567:600–608, 1991.

[8] M. Irani and S. Peleg. Improving resolution by image registration. *Graphical Models and Image Processing*, 53:231–239, 1991.

[9] S. P. Kim, N. K. Bose, and H. M. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Trans. ASSP*, 38:1013–1027, 1990.

[10] Y. Yacoob L. Davis, R. Chellappa and Q. Zheng. Visual surveillance and monitoring of human and vehicular activity. In *Proc. IUW*, pages 19–23, 1997.

[11] R. E. A. C. Paley and N. Wiener. *Fourier Transforms in the Complex Domain, Colloquium Publications, Vol. 19.* Am. Math. Soc., 1934.

[12] A. Papoulis. A new algorithm in spectral analysis and bandlimited signal extrapolation. *IEEE Trans. Circuits and Systems*, 22:735–742, 1975.

[13] A. Papoulis. Generalized sampling expansion. *IEEE Trans. Circuits and Systems*, 24:652-654 1977.

[14] A. Papoulis. *Signal Analysis.* McGraw-Hill, 1977.

[15] M. S. Scivier and M. A. Fiddy. Phase ambiguities and the zeros of multidimensional band-limited functions. *J. Opt. Soc. Am.*, 2:693–697, 1985.

[16] H. Shekarforoush. *Super-Resolution in Computer Vision.* PhD thesis, University of Nice, 1996.

[17] H. Shekarforoush, M. Berthod, and J. Zerubia. 3D super-resolution using generalized sampling expansion. In *Proc. ICIP*, Vol. 2, pages 300-303, 1995.

[18] H. Shekarforoush, M. Berthod, and J. Zerubia. Subpixel image registration by estimating the polyphase decomposition of cross power spectrum. In *Proc. CVPR*, pages 532–537, 1996.

[19] H. Shekarforoush, M. Berthod, J. Zerubia, and M. Werman. Sub-pixel Bayesian estimation of albedo and height. *IJCV*, 19:289–300, 1996.

[20] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty IV: Extension to many dimensions; generalized prolate spheroidal wave functions. *Bell System Tech. J.*, 43:3009–3057, 1964.

[21] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty V: The discrete case. *Bell System Tech. J.*, 57:1371–1430, 1978.

[22] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty I. *Bell System Tech. J.*, 40:43–63, 1961.

[23] S. Srinivasan and R. Chellappa. An integrated approach to image stabilization, mosaicking and super-resolution. In *Proc. IUW*, pages 247–253, 1997.

Appendix

Proof of Theorem 2.2: First, a well-known result is applied by change of variables to polar coordinates:

$$\begin{aligned} u - u_k &= \omega cos(\varphi) \qquad & v - v_k &= \omega sin(\varphi) \\ x &= tcos(\theta) \qquad & y &= tsin(\theta) \end{aligned} \tag{21}$$

This leads to the following representation of the Fourier inversion theorem for $\hat{h}$:

$$h_k(t)exp(-it\rho_k) = \int_0^\infty \omega \hat{h}_k(\omega) J_0(t\omega) d\omega \tag{22}$$

where $\rho_k = u_k cos(\theta) + v_k sin(\theta)$, $\tilde{h}_k(\omega) = h_k(u,v)$ stands for the Hankel transform of $h_k(t) = h_k(x,y)$, and J_0 is the zero-order Bessel function of the first kind.

Applying the Hankel inversion theorem we then obtain

$$\tilde{h}_k(\omega) = \int_0^{b_k} t h_k(t) exp(-it\rho_k) J_0(t\omega) dt \tag{23}$$

where b_k is the radius of local isotropy.

Since at $(u,v) = (u_k, v_k)$ we have $\tilde{h}_k(\omega) = 0$, we can write

$$\int_0^{b_k} t h_k(t) exp(it\rho_k) dt = 0 \tag{24}$$

Substituting from (22) into this last integral equation, we get

$$\int_0^{b_k} t \int_0^\infty \omega \tilde{h}_k(\omega) J_0(t\omega) d\omega dt = 0 \tag{25}$$

which, after changing the order of integration, yields

$$\int_0^\infty \omega \tilde{h}_k(\omega) \int_0^{b_k} t J_0(t\omega) dt d\omega = 0 \tag{26}$$

Using the well-known identity

$$\int_0^a t J_0(t) dt = t J_1(at) \tag{27}$$

we deduce that

$$\int_0^{b_k} t J_0(t\omega) dt = \frac{b_k J_1(b_k \omega)}{\omega} \tag{28}$$

Therefore

$$\int_0^\infty b_k \tilde{h}_k(\omega) J_1(b_k \omega) d\omega = 0 \tag{29}$$

Finally, using the following integral relation for Bessel functions of the first kind [4]:

$$\int_0^\infty J_n(b\omega)exp(\pm i\omega a) = \frac{i^{\pm(n+1)}}{\sqrt{a^2-b^2}}\left(a-\sqrt{a^2-b^2}\right)^n \tag{30}$$

we conclude that

$$\lim_{b_k\to 0} b_k\hat{h}_k(\omega) = exp(-i\omega t_k) \tag{31}$$

where $t_k = (x_k^2 + y_k^2)^{\frac{1}{2}}$ and (x_k, y_k) is a constant vector in $\mathbf{T}$.

Errata

1. Page 7 paragraph 5 immediately after equation (15) should read:

where $\hat{h}_{g1} = \hat{h}_1 + \hat{n}_1 \frac{\hat{f}_1^*}{|\hat{f}|}$ and $\hat{h}_{g2} = \hat{h}_2 + \hat{n}_2 \frac{\hat{f}_2^*}{|\hat{f}|}$ with * denoting the complex conjugate and $|\hat{f}_1| = |\hat{f}_2| = |\hat{f}|$.

2. Page 7 equation (16) should read:

$$| \hat{g}_1 \hat{g}_2^* | = c \, | \hat{h}_{g1} \hat{h}_{g2}^* | \qquad \text{where } c = |\hat{f}|$$

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1998	Technical Report

4. TITLE AND SUBTITLE

Data-driven Multi-channel Super-resolution with Application to Video Sequences

5. FUNDING NUMBERS

N00014-95-1-0521

6. AUTHOR(S)

H. Shekarforoush and R. Chellappa

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Center for Automation Research
University of Maryland
College Park, MD 20742-3275

8. PERFORMING ORGANIZATION REPORT NUMBER

CAR-TR-885
CS-TR-3898

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street, Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release.
Distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A method is proposed for super-resolving multi-channel data with applications to PREDATOR video sequences. Using a generalization of Papoulis' sampling theorem, a closed-form solution has been obtained leading to a high-speed algorithm which can be realistically applied to large data sets such as video sequences. In existing multi-frame methods it is a common practice to assume that the channel transfer functions are known and invariant from one frame to another, using empirical models such as Gaussian, sinc, etc. We have assumed that the transfer functions are unknown and may vary even when the same sensor is employed, and hence use the observed data to derive the Point Spread Function (PSF) for each frame. The estimated PSFs are used in the super-resolution algorithm. Results on PREDATOR video images are then given.

14. SUBJECT TERMS

Super-resolution, video data processing, multi-channel sampling, data-driven estimation

15. NUMBER OF PAGES

23

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.